# All for one, one for all

*A look at the state of the art in*

## SERVER CLUSTERING.

BY MARC SONGINI

Sometimes there is a vast valley between the mounds of hype vendors build over a technology and the rolling, sometimes craggy hills of reality where users work.

Such is the case with most of the latest hype surrounding clustering.

Clustering, in a nutshell, is the ability to strap together multiple CPUs — be they PC, mid-range or mainframe — into a single, ideally more powerful entity than a stand-alone box. Vendors tout clustering as a way to boost scalability, performance and availability.

For years, vendors such as Digital, IBM, and Sun have been tying CPUs together via software and hardware techniques to boost performance and provide failover functions. The efforts have met with varying degrees of success. IBM this summer announced clustering capabilities for its RISC System/6000 AIX workstations and AS/400e mid-range family.

"Digital invented clustering back in 1983," says Jackie Clement, an AlphaServer product manager at Compaq's Digital subsidiary. The technology has thousands of connected nodes in its installed base. For reliability, clustering offers "no single point of failure," she says. "Any component can fail and the system can keep running. Everything is hot-swappable."

While hardware vendors continue to expand their cluster offerings, few of them have generated the clamor Microsoft has with its Microsoft Cluster Server product, popularly known as Wolfpack. Wolfpack is the technology Microsoft says will provide server failover protection and multiple server clustering capabilities. But most of the promise of Wolfpack remains just that — a promise.

According to Microsoft, most users are content with the simple failover and availability features the company already offers. The software giant denies it has welshed on any of its promises.

"We've been developing clustering in two phases," says Mark Hassall, product manager for Windows NT Server at Microsoft. The first stage is availability, Hassall says, and the second is scalability. Microsoft plans to have a two-node-plus cluster in beta after the release of Windows NT 5.0, which does not yet have a release date.

Hassall says clustering does not necessarily mean linear scalability. "Scalability depends on the hardware and having cluster-aware applications. A lot of things have to happen to deliver scalable clusters. Customers will have to wait some time."

Another hotly hyped clustering technology is Red Hat's Extreme Linux Version 1.0. Extreme Linux runs on a variety of hardware boxes and can, in theory, link an unlimited number of CPUs. The product links Red Hat's popular Linux operating system with add-on clustering technology, code-named Beowulf, which was developed at NASA's Ames lab. The combination of Red Hat's Linux and Beowulf clustering, while promising, is more hype than reality. That is mainly because there are few business applications that can take advantage of the system. Linux also requires users to do much of the setup and configuration — not a pleasant idea for many customers.

Oracle has tried to get into the clustering game as well. The company's Parallel Server addresses failover and parallel computing, but Parallel Server currently is limited to six nodes.

Vendor delivery issues aside, analysts say clustering may just be too complicated for users to get their hands around.

"Customers are concerned that clustering is going to be too complicated to install or be something they don't need because it is so complicated to get up and running," says James Gruener, an analyst at Boston-based Aberdeen Group. "It hasn't taken off in the industry as fast as a lot industry watchers said it would."

"Wolfpack got a lot of press, but it hasn't delivered very much," says David Floyer, an analyst at the consultancy International Data Corp. in Framingham, Mass. The reality is that the triple promises of clustering — availability, linear scalability and a reduction in costs — have only been partially realized, if it all.

Clustering generally works when used to do data warehousing or to offer some scalability. Anything more complex, and clustering doesn't cut it, Floyer says. The technology generally succeeds when it's used in simple operations, such as supporting banks of automatic teller machines, he adds.

At least one user agrees with that assessment.

"Clustering requires too much manpower," says Stephen Lewis, senior network design engineer for Signal Corp., a Fairfax, Va.-based IT integrator. "Clustering requires more administration than it's worth," he says. Tying together servers with a Fibre Channel pipe may be a better option than clustering, Lewis says.

When two servers connected by Fibre Channel are passing data back and forth, they work so fast no back-up operations are needed to duplicate data. If one server crashes, it doesn't matter, because it will have already replicated its data on the other server before it goes down.

With the increased use of high-speed Fibre Channel connections and other technologies, clustering is becoming more viable, Floyer says. With Fibre Channel, users can more easily make connections among every disk and server in a cluster system. In addition, there will be new management tools added to cluster offerings that will make them more usable in the future, Floyer says. Costs should start dropping, as well. ●

### Who started the buzz?

### CLUSTERS

**Digital**

**Where is it now?**

Digital's VMS clusters were first delivered with VMS 4.0 in 1984. The inception of clusters was a collaborative effort of many people at Digital. Key players included Dick Hustvedt, Dave Thiel, Gordon Bell and Bill Strecker. Many others from the VMS Engineering organization contributed as well.

**Market size:**

There are some 150,000 VMS clusters today, Digital executives say. An average cluster includes three systems. OpenVMS clusters document support for 96 systems per cluster, but Digital claims to have customers who have linked as many as 150 systems in one cluster. Market research from International Data Corp. says the clustering software market is worth between $70 million and $80 million annually.

Clusters are considered by many experts to be Digital's most significant contribution to the industry.

**"Today, virtually every major computer system vendor offers some type of system called a cluster. Digital's technology is still the high bar against which other clustering schemes are measured."**

*— Digital executives*